



TITLE:

漢籍の電子的な翻刻について：『説文解字注』の Unicode 轉寫を事例として

AUTHOR(S):

白須, 裕之

CITATION:

白須, 裕之. 漢籍の電子的な翻刻について：『説文解字注』の Unicode 轉寫を事例として. 東方學報 2014, 89: 238-211

ISSUE DATE:

2014-12-20

URL:

<https://doi.org/10.14989/196357>

RIGHT:

漢籍の電子的な翻刻について

——『説文解字注』の Unicode 轉寫を事例として——

白 須 裕 之

目 次

- | | |
|------------------|--------------|
| 1 はじめに | 6 形式化について |
| 2 翻刻の方法 | 7 おわりに |
| 3 Unihan データベース | 附録A 「字畫原始」の例 |
| 4 事例 | 附録B 「賣」と「賣」 |
| 5 文字コードによる轉寫を越えて | 附録C 「昏聲」の字 |

1 は じ め に

古典籍の電子的な表現には、主に畫像を利用するものと文字コードを利用するものの二つがある。ここでは文字コードを前提に漢籍の電子的な翻刻を考え、その問題點を確認し、將來の文字の電子的な表現への布石とする。文字という對象は概念的な存在であり、解釋や知識による揺れをもつ。漢籍の電子的な翻刻には、このような歴史的なものを對象として同定する方法と、それを指し示す方法が必要である。本稿では『説文解字注』を事例として、電子的な翻刻の實際の問題點を提出する。

『説文解字注』（以下、段注）の讀書會である「點注會」では、段注のより良い電子テキストの構築を目指し、講讀・翻刻・標點など、段注の様々な面について検討している¹⁾。本稿はその會での経験と議論に基づいて、漢籍の電子的な翻刻の實際と、そこから得た知見について報告する。ここでは Unicode 文字集合による轉寫について述べ、その問題點を議論し、將來の漢籍の電子的な翻刻の姿を描く。

1) 「點注會」は 2011 年に京都大學人文科學研究所准教授古勝隆一氏と筆者が發起人になり、京都大學文學部中國哲學史專修の有志の皆さんとで構成されている。

現在の電子テキストやテキストデータベースが採用している文字の電子化の方法は、符號化文字集合を用いるものが一般的である。その符號化文字集合の中でも、漢字文化圏における古典文獻（漢籍）のテキストを扱うには、その文字集合の普及の廣さと、收録されている漢字の多さという點において、Unicode で定義されている符號化文字集合が、殆ど唯一の選擇肢になっている²⁾。

Unicode コンソーシアムでは、Unicode の符號化文字集合における漢字が持つ各種情報を提供するために、Unicode Han（漢）Database（Unihan データベース）を公開している。Unihan データベースでは、Unicode に收録されている漢字の標準字形、收録の際に利用した典據、典據となった様々な文字集合へのマッピング、標準的な字典への参照などの情報を Web で提供している。

Unicode の文字集合による漢籍の電子的な翻刻には、この Unihan データベースを参照する必要がある。Unihan データベースの詳細については、節 3 で見ることにする。

2 翻刻の方法

文獻 [10] において、「情報の流れ」理論の立場から文字概念の電子的な表現を検討した。また、その應用として、古典籍の翻刻に對する「文字轉寫の理論」を提出した。本稿の符號化文字集合による翻刻も、この「文字轉寫の理論」の考えに基づく³⁾。ここではその概略について述べる。詳しくは文獻 [10] を参照していただきたい。

文字コードによる翻刻では、文獻に對する解釋を経て、釋文（解釋文字列）が構成され、その釋文が符號化される。その際の符號化には符號化文字集合の文字概念が用いられる。

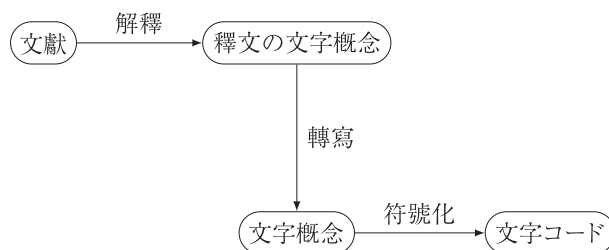


圖 1 釋文と轉寫

2) JIS 漢字を利用した翻刻については、文獻 [13] を参照していただきたい。

3) 本稿では文獻 [10] の形式化を直接には利用しない。形式化のその後の發展については文獻 [11] を参照していただきたい。

その様子を図 1 で示した。転寫には釋文の文字概念と、それを符號化するための文字コードの文字概念という二種の文字概念が必要である。

電子化の際に問題となるのは、文獻の文字概念（釋文の解釋）と文字コードの文字概念が形式化されていない點である。本稿の立場では、文字コードの文字概念は Unihan が提供する標準字書への参照によって決まるものとする。また、文獻の轉寫によって表現された電子テキストは、そのテキスト自身が文字コードの文字概念への関係を表現している譯であり、文獻の文字概念（釋文の解釋）は轉寫自身が表現しているとして扱う。

これは従來の翻刻の作業となんら變わる所はないが、「文字轉寫の理論」の考えを意識した翻刻は、文獻 [10] [11] の形式化に對して實際の文字データを提供することになる。更にその作業で得られた事例を検討することによって、文字表現のみではなく、電子テキストの新たな表現の可能性も模索することができると考える。

2.1 翻刻の方針

Unicode の文字集合を使用して、段注の電子テキストを作成する上の方針について述べる。

文獻 [5] では、『説文』の（親字である）篆文を Unicode 文字集合に照應するための Unicode Best Match という基準が提出されている。まず一つの篆文に對應する可能性のある文字の候補が Unicode 文字集合から選らばれて、それを列舉した表 “The SWJZZUnicode 4.0 Variant Mapping” を作成する。この表には Unicode の標準字形へのフィードバックも考慮して、照應の判定結果も保持されている。判定には標準字形と説解の部分字形、『漢語大字典』の參照先が利用されている。

本稿では篆文を楷書にマッピング出来れば良い譯ではなく、注の文字も含めて翻刻を行なう。文字を選ぶ基準は Unicode Best Match と同様の考えを用いるが、候補の列舉には後で述べる「標準字書」の記述を重視し、標準字形には重きを置かない。この理由については事例で説明する。このような翻刻の方針によって作成された電子テキストでは、個々の文字コードが何故にそのように轉寫されたのかを、「標準字書」を通して確認できる。

段注の翻刻では、段氏が使用する文字を通して段氏の正字意識に注目し、その文字概念を明確化することが目標である。文獻の編集では、解釋と記録を整然と分けることが前提とされる。畫像を利用する場合がこれに當るであろう。上で述べたように、翻刻ではテキストの読みや解釋が釋文の作成に本質的な役割を果している譯であり、ここでの翻刻は解釋をどこまで形式化できるか、どのような情報化が可能であるかを研究するのが我々の使命であると考えている。

3 Unihan データベース

本節では Unihan データベースがどのようなデータを提供していて、このデータベースを漢籍の電子的な翻刻に如何に利用するかを説明するために、Unihan データベースの概要について纏めておく。

3.1 Unihan 文字集合

Unihan のデータモデルを理解するための前提として、Unicode において定義された概念 character, property について纏めておく⁴⁾。Unicode における文字の扱いは、文献 [6] の節 2.2 Unicode Design Principles における 10 の原則に端的に表現されている。そこで述べられているように、全ての原則を同時に充すことは必ずしも可能ではない。従って、シンプルさ、効率、既存の規格との互換性とのバランスを考慮して、設計が行われている。以下、Unihan を理解する上で、関係がある原則について見ていく。

1. Character, not glyphs : The Unicode Standard encodes character, not glyphs.
2. Semantics : Characters have well-defined semantics.
3. Unification : The Unicode Standard unifies duplicate characters within scripts across languages.

1 によって符號化の対象は character であることが分かる。その character は、以下に述べられているように、抽象的な表現であり、glyph とは厳密に區別される概念である。2 によって character は明確な semantics を持ち、その semantics は property によって與えられる。3 により Unicode では各言語を越えて文字を統合化して、文字集合を規格化する。

漢字の統合化を説明するために、漢字の性質をモデル化した三次元概念モデル（文献 [6] 417 頁）がある。漢字についての以下の三つの属性が各々の軸を構成する。

- ・ X 軸：セマンティック（意義、機能等）
- ・ Y 軸：抽象的形狀 abstract shape
- ・ Z 軸：具體的形狀 actual shape（具體的な書體 type face）

4) 用語 character, property は Unicode 固有の概念であるため、本稿では譯語を用いることなく、そのまま使用する。

Unicode の規格において、この三次元概念モデルに実際に漢字を配置するのは、統合規則（Unification Rule）と抽象的形を決める規則の二つである。UniHan データベースを現代の文献の電子化に利用する場合には、この二つの規則は有用であるが、後に述べるように古典籍の電子化にはそれほど有用でない⁵⁾。

そこで Unicode における漢字を扱うために、個々の文字に対する property のみを対象とする⁶⁾。以下に property によって構成される「UniHan データモデル」について説明する。

3.2 UniHan データモデル

図 2 は UniHan データベースで扱われるデータをクラス図で表現したものである⁷⁾。クラス Ideograph は漢字を表現するクラスで、これには属性 codepoint が付随している。UniHan データベースの property は、各漢字にプロパティの名前とそれに對應する値を指定するという形式が基本である。図ではこのプロパティの持つデータを、クラス Property とクラス Field の二つで表現している。クラス Property は property の値を、クラス Field は property の名前や管理情報を持っている。これはクラス Property が二つのクラス Ideograph と Field のある種の関連クラスになっていることを示している。また Field を種類ごとに分類するのが、メタタイプとしてのクラス FieldType である。

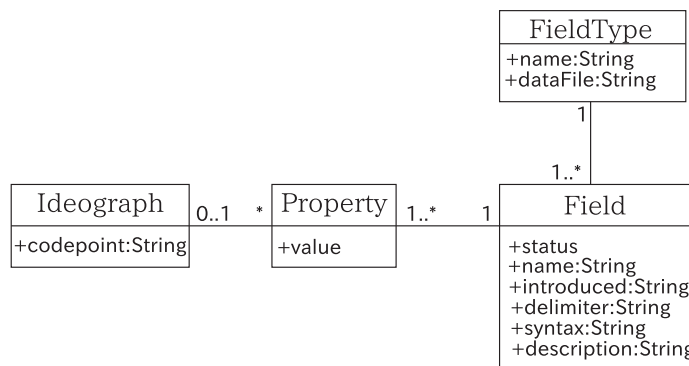


図 2 UniHan データモデル

5) 包攝標準の擴張による異體處理によって、電子的な翻刻の精度を上げる研究には文献 [12] などがある。本稿ではそのような方策は取らない。古籍の場合では、包攝標準によって同定された文字が形の再現性を高くできても、標準字書の記述に矛盾することが往々にして見受けられるからである。

6) 各 property 値の読み方や意味について、詳しくは文献 [6] を参照していただきたい。

7) クラス図については UML の入門書、例えば文献 [7]などを参照していただきたい。

Unihan データモデルをこのように解釈すると、文字概念を表現する方法として、文字のセマンティックを名前付きプロパティの集まりによって指定していることが一目瞭然である。Unihan はこのプロパティの規格を決めていることになる。

3.3 標準字書について

Unihan データモデルの property の中で本稿で特に重要視するのが、kIRGHanyuDaZidian, kSBGY, kIRGKangXi の三つである。これらのプロパティ値は Unihan が決めた標準字書『漢語大字典』『宋本廣韻』『康熙字典』への参照位置（頁と頁内番號）を示す。以下に例を使って各値の読み方を簡単に示す。各字書のバージョンや詳細については文献 [8] を参照していただきたい⁸⁾。

- ・『漢語大字典』kIRGHanyuDaZidian 32264.080: 3 卷 2264 頁の親字 8 文字目,
- ・『宋本廣韻』kSBGY 364.38: 364 頁の親字 38 文字目,
- ・『康熙字典』kIRGKangXi 1187.060: 1186 頁の親字 6 文字目.

本稿での翻刻には標準字書の情報を用いるため、各文字 Ideograph に對するこれらの値を調べる必要がある。現在、Unihan で提供している文字を探すためのツールとしては、Unicode コンソーシアムが提供する以下の二つの Web 検索システムがある。

- ・“Unihan Database Lookup”: コードポイント及び文字のエンコードから漢字情報のページを検索する⁹⁾,
- ・“Search the Unihan Database”: 漢字の簡単な意味、及び各種の音から文字を検索する¹⁰⁾.

この他、Unihan の漢字情報ページを検索するためのツールには様々なものが存在すると思われるが、ここでは述べない。

8) kIRGHanyuDaZidian, kIRGKangXi の数値の最後一桁（小数点 3 桁目）が 0 のときは、その文字が字書のその位置に実際にあることを示す。それ以外の場合は整列のために利用するので、本稿では用いない。

9) <http://unicode.org/charts/unihan.html>

10) <http://unicode.org/charts/unihansearch.html>

4 事 例

以下では、段注での参照箇所を示すために、「段注三篇下 28b」のような略號を使用する。これは『説文解字注』三篇下の 28 葉裏を示す。a が表、b が裏である¹¹⁾。なお斷わらない限り、版本は「經韻樓本」¹²⁾を用いる。

ここで扱う事例は中國古典學では常識に屬するものであろうが、轉寫の作業者の参考に供するとともに、中國古典學者にも Unihan の使い方の参考になればと考えている。

4.1 形の似ている字の區別

まず字形の似ている二つの文字「几」と「兀」が、『説文』ではどのように扱われているかを見てみる。その後、段注でこれらに似た文字が現れたときに、どの文字コードを使用すべきかの判断を通して、Unihan の使用例を見て行く¹³⁾。

段注の三篇下几部と十四篇上兀部の部首字には、それぞれ以下のように記されている。

兀 (几) 鳥之短羽飛几几也。象形。凡几之屬皆从几。讀若殊。市朱切。按以彡從几聲求之。古音在四部。(段注三篇下 28b)

兀 (几) 尻几也。尻各本作踞。今正。尻几者、謂人所尻之几也。尻、処也。処、止也。古之尻今悉改爲居。... 象形。象其高而上平可倚。下有足。居履切。十五部。周禮五几。玉几、彫几、彤几、鬋几、絜几。周禮司几筵職文。鬋几、今周禮作漆几。蓋許所據不同。鬋者、漆也。凡几之屬皆从几。(段注十四篇上 28a)

Unihan データベースにおいて、「几」U+20627 と「兀」U+51E0 のそれぞれの文字コードが指している標準字書の参照先は以下のようにになっている。

-
- 11) その他の段注に關する用語等は文獻 [15] に準ずる。
 12) 本稿で用いる段注の畫像は早稻田大學古典籍總合データベース <http://www.wul.waseda.ac.jp/kotenseki/> (請求記號ホ 04 00026 0001) からのものである。
 13) 事例説明のために『説文』や段注の轉寫した記述を屢々用いる。使われる轉寫はその後の議論を前提としているので、その點には注意していただきたい。また、組版の都合で篆文には SW フォントを使用しているので、精確を期すためには段注の經韻樓本を参照いただきたい。

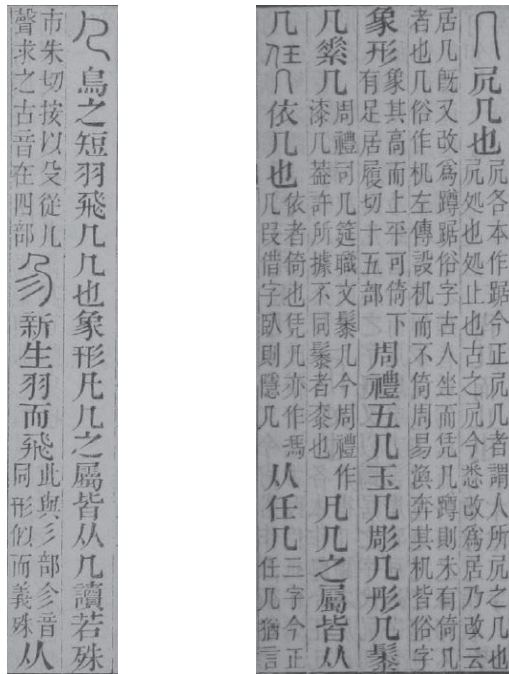


圖3 「几」(三篇下 28b), 「几」(十四篇上 28a)

・「几」U+20627

- 漢語大字典 (10276.010)《說文》“鳥之短羽飛几几也。象形”...《廣韻》市朱切, 平虞禪。侯部。
- 康熙字典 (0134.010) 134 頁 1 文字目 [廣韻] 市朱切...
- 廣韻 (076.12) 上平聲十虞小韻「殊」市朱切

・「几」U+51E0

- 漢語大字典 (10275.190)《說文》“几踞, 几也。象形”...《廣韻》居履切, 上旨見。脂部。
- 康熙字典 (0133.570) [唐韻] 居履切...
- 廣韻 (247.32) 上聲五旨小韻「几」居履切

各字典への参照には矛盾はないので、三篇下「几」と十四篇上「几」の字は、それぞれ U+20627 と U+51E0 で轉寫して差し支えないことが分かる。『康熙字典』では二字の字體の區別は最後の跳ねの有無で區別するが、『廣韻』では最後の跳ねの有無で區別するよりも、段注と同様に左角の縦が突き出ているかどうかで區別し、突き出ているのが三篇下「几」である。Unicode の區別は『康熙字典』の區別を踏襲している。

以上、本稿での Unicode 轉寫は Unihan の標準字書の記述と照らし、形音義を総合的に

判断して決めることになる。

4.2 字形に依れない例

まず図4の画像，段注二篇下8b 辵部「避」字の部分を見てほしい。説解「回也」の注「上文回辟之回訓袞」の次の字をどのように Unicode 転寫すべきであろうか。

「避」の注に使ってある字形からすると，「𨾏」U+2018E に転寫して良いようにみえるが，これは八篇上衣部の「褌」U+465F（または「𨾏」U+27755）の異體字で，ここで検討している字とは別字である¹⁴⁾。

「訓袞」に注目して，八篇上 64a 衣部「袞」を見ると以下のように書かれている。

𨾏 袞也。交部曰。袞者。袞也。二篆爲互訓。小徐本作紕也。非是。袞今字作回。
袞今字作邪。毛詩傳曰。回，邪也。从衣。牙聲。似嗟切。古音在五部。（八篇上 64a）

また，十篇下 9b 交部「𨾏」を見ると以下のように書かれている。

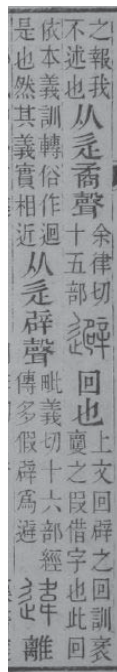


図4 「避」(二篇下 8b)

14) Unihan での「𨾏」U+2018E が指し示す『漢語大字典』の参照先には，「同“𨾏（褌）”。」とある。

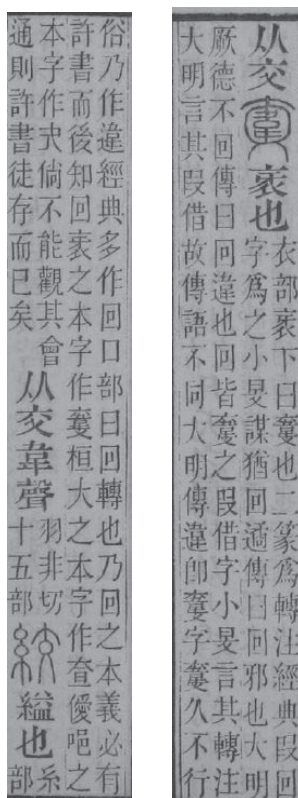


図5 「𦐇」(十篇下9b)

𦐇

衰也。衣部衰下曰。𦐇也。二篆爲轉注。經典段回字爲之。小旻。謀猶回適。傳曰。回，邪也。大明。厥德不回。傳曰。回，違也。回皆𦐇之段借字。... 从交。韋聲。羽非切。十五部。(十篇下9b)

衣部「衰」と交部「𦐇」は互訓（轉注）であり，「回」は「𦐇」の段借字であるので，現在問題にしている「避」注の字は，交部「𦐇」であることが判明する¹⁵⁾。それではこの字をどのように轉寫すべきであるか。Unicodeによる轉寫としては以下の候補があげられる。

・「𦐇」U+293AE¹⁶⁾

15) 例によって，この文章は以下の論述に従った轉寫を利用している。

16) 『漢語大字典』10293.130, 74506.050

・「𡗗」U+293C9¹⁷⁾

直前の小節と同じように Unihan のそれぞれの参照先を見ると、これら二字「𡗗」「𡗗」は異體字であり、交部の同一篆文の異なる楷化字である。ここでは交部「𡗗」の注で段氏が使っている字形に近い「𡗗」字を、「避」字の注の轉寫に使用することにする。

(避) 回也。上文回辟之回訓衰。𡗗之段借字也。此回依本義訓轉。俗作迴是也。然其義實相近。从辵。辟聲。毗義切。十六部。經傳多假辟爲避。(二篇下 8b)

文献 [5] では、誤って交部「𡗗」を「𡗗」に轉寫している。篆文の字形に注目したために、Unicode best match の候補に「𡗗」を含めたことが原因と思われる。

4.3 音注に依る例

我々が良く使用する字形の爲に、その諧聲符の字を間違えやすい字に「呈」「廷」などの「壬聲」の字がある。「呈」「廷」の部分字形にしばしば「壬」が使われることがあるためである。「壬聲」は七部、「壬聲」は十一部である。たとえば、段注の二篇上 21a 口部「呈」を見ると、その聲符がどちらであるか字形だけでは判別しづらい。

(呈) 平也。今義云示也。見也。从口。壬聲。直貞切。十一部。壬之言挺也。故訓平。(二篇上 21a)

そこで音注の「呈」が十一部であることを利用すれば「壬聲」であることが判明する¹⁸⁾。

4.4 Unihan の矛盾

Unihan の参照に矛盾がある場合を取り上げる。段注の二篇上 9a 牛部「𡗗」の記述は以下である。

(𡗗) 兩壁耕也。壁當作辟。辟是旁側之語。莊廿一年左傳。鄭伯享王子闕西辟。服虔云。西辟，西偏也。兩辟耕謂一田中兩牛耕。一從東往，一從西來也。此耕字自人牛言之。與木部六又犁自器言之不同。从牛。非聲。此形聲包會意。非从飛下𡗗，取

17) 『漢語大字典』74507.120; 『康熙字典』1394.360

18) これらの辨別には現代音を知っていることが助けになる場合である。

其相背。非尾切。十五部。廣韵入去聲。一曰覆耕種也。此別一義。未聞。讀若
匪。(二篇上 9a)

この注にある「莊廿一年左傳」の「廿」字の轉寫を考える。Unihan の「廿」U+5344
の參照先にはそれぞれ以下の説明がある。

- ・『漢語大字典』10058.100 十部「廿」：同“廿”。《玉篇・十部》：“廿，二十并也。今直為二十字。”
- ・『宋本廣韻』038.07 上平聲三鍾小韻「恭」九容切：「竦手也說文本居竦切」

『廣韻』の記述は『説文』三篇上収部の部首字「収」(段注三篇上 35b)であり、『漢語大字典』とは別の字を参照している。『廣韻』には「廿」字に似た「廿」の意味の字はなく、この文字に関して Unihan は『廣韻』の上記の參照先を示すべきではなかったことになる。因みに Unihan には「廿」U+5EFE があり、どちらの字典の參照も『説文』の「竦手也」の字を指している。

Unihan ではこの例のように標準字書の記述のあいだに矛盾を生ずる字が多く存在する。轉寫にはその時々で総合的な判断が必要になる。

5 文字コードによる轉寫を越えて

前節では Unicode 文字集合による轉寫の事例を見てきた。本節では文字集合による轉寫の枠組みからはみ出すような事例であり、かつ文字概念の形式化を検討する上で参照すべき事例について見る。

5.1 「𠂔」と「𠂔」

日本語用法で「𠂔」(常用漢字)は「レイ、もど・す、もど・る」として使用し、「𠂔」はその舊字として扱う¹⁹⁾。すなわち、二つの字は異體字として扱われている。一方、『説文』(段注を含む)では以下のように音も意味を異なる別字である。

𠂔 (𠂔) 輻車蜀推戶也。从戶大聲。讀與欽同。徒蓋切(『説文』)

19) 例えば、文獻 [9] を参照。

十二篇上戸部)

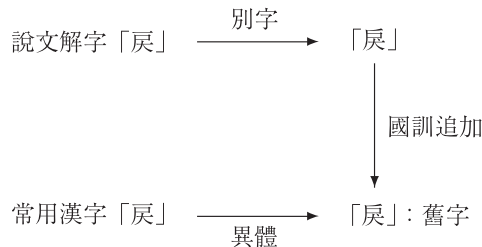
𠂔

(𠂔) 曲也。从犬出戸下。𠂔者，身曲𠂔也。郎計切 (『説文』

十篇上犬部)

個々の文字は、字書や文献の解釈によって得られる文字集合の中で考えられる存在である。この文字間の張り合いが文字體系を構成する。また、それぞれの體系どうしは無関係という譯ではなく、傳達の手段として使用された文字に歴史が積み重なり、異なる體系に属す文字にも歴史的な連鎖が備わっていく。

この場合の例を単純化してしまうと、『説文』の「𠂔」に國訓が加わり、それを常用漢字では「𠂔」と表記するようになる。Unicodeのようなセマンティックを持つ文字集合では、説文の「𠂔」と常用漢字の「𠂔」には同一の文字コードを使用することになるが、漢字情報の形式化では異なる文字體系（世界）の文字として解釋する必要がある。



また「別字」「異體字」という文字間の関係は文字體系（世界）に依存した事実であり、文字體系そのものは「世界の可能的な状態」として解釋することができる。このように文字を體系の中で捉える發想については節6で詳しく見る。

5.2 「賣」と「賣」

ここでは部分字形などで屢々混同して使用されている「賣」と「賣」について見る。「賣」は大徐本 [2] の『説文』六篇下出部に以下のようにある。

𠂔

(賣) 出物貨也。从出从買。莫遯切

一方、「賣」『説文』六篇下貝部に以下のようにある。

𠂔

(賣) 銜也。从貝畜聲。畜，古文陸。讀若育。余六切

これらの二字について段注では以下のように述べられている。

(賣) 出物貨也。《周禮》多言賣儋，謂賣買也。从出。从買。出買者，出而與人買之也。《韻會》作買聲，則以形聲包會意也。莫遯切。十六部。(六篇下 2b)

(賣) 衍也。衍，行且賣也。賣字不見經傳。《周禮》多言儋。儋訓買。亦訓賣。《胥師》飾行儋慝。《賈師》貴儋者。蓋即說文之賣字。而《說文》人部「儋，見也」。則今之觀字也。《玉篇》云：「賣或作粥鬻」。是賣，鬻爲古今字矣。按賣籀變作賣。易與賣相混。从貝。畚聲。畚古文陸。見貝部。讀若育。余六切。三部。(六篇下 21b)

段注ではこれら二字を異なる古音分部に配置する。音韻情報を纏めておくと、以下のようになる。

1. 「賣」：出部，莫遯切。十六部。
2. 「賣」：貝部，余六切。三部。(「育」の音)

5.2.1 諧聲符と部分字形

今、「賣」を諧聲符として持つ字を考える。例えば「讀」「牘」「瀆」「續」「贖」は全て古音三部に屬す。これら五字に對する字形して、旁を部分字形「賣」とするものを屢々見かける。形は似ていても、『說文』から見たとき、「賣」ではなく「賣」を部分字とするのが適切である。詳しくは附録 B を見られたい。

5.2.2 『慧琳音義』の例

以下は『慧琳音義』の音と字が矛盾している例である。『大正新脩大藏經』の慧琳『一切經音義』卷第八十二「大唐西域記」第二卷から引用する²⁰⁾。

鬻賣 上融宿反，俗字也。說文正體作賣，衍也，賣也。從貝畚聲，畚，古文陸字也。贖字從賣從貝，從賣者俗字也。下埋解反，從出從買，從土者俗字也 (T2128_ 54.0838c24~25)

「說文正體作賣，衍也，賣也。」という箇所では、「鬻」が反切「融宿反」を持つ事及び説解「衍也」より、「賣」は「賣」であるべきである。また、「贖」字從賣從貝，從賣者俗

20) 「SAT 大正新脩大藏經テキストデータベース 2012 版」<http://21dzk.l.u-tokyo.ac.jp/SAT/ddb-bdk-sat2.php> より引用，但し Unicode 轉寫は著者による。

字也。」の箇所も「賣」は「賣」とすべきである²¹⁾。これらの事実は音韻情報から機械的に推論できる事柄である。寫本や版本を校訂するときと同様に、計算機による音韻情報の提供は電子的な翻刻の精度を上げることに繋がるであろう²²⁾。

5.3 段氏の特殊な用字の例

ここでは段玉裁の使用している特殊な用字の例について述べる。まず圖6を見てほしい。この一篇下21bの艸部「𦵏」の注において、部分字形「舌」の部分に「ノ」と「右」を合字したような字形が使われている。「𦵏樓」「𦵏果」「古活切」の箇所に見える。以下、この特殊な形の「舌」字を「ノ右」で示すことにする。

他にも部分字形「ノ右」を使用する字は以下に見える。

- ・四篇上3b 目部「睨」：「刮」字
- ・八篇上19a 人部「𦵏」：「𦵏」「活」字
- ・十篇下41a 心部「𦵏」：「𦵏」「𦵏」字

楷書「舌」には來源として「昏」と「舌」の二つがある。『説文』に見える「昏聲」の字は、「話」「活」「括」など十八字を数える。現代の楷書ではこれらの字の部分字形を「舌」とするのが普通である。その内譯は附録Cを参照していただきたい。段氏は二篇上26bの「昏」で「凡昏聲字𦵏變皆爲舌。」と述べている。



(昏) 塞口也。廣雅釋詁曰。昏，塞也。易六二。括囊无咎。括即昏字也。

从口。𦵏省聲。𦵏即氏部𦵏字。𦵏變或作𦵏。或作𦵏。凡昏聲字𦵏變皆爲舌。如括刮之類。古活切。十五部。(二篇上26b)

『説文』の「𦵏」「𦵏」(昏聲)の二字は「𦵏」「𦵏」(舌聲)とは別字であり、この二字に部分字形として「舌」を使うことは混亂を招くおそれがある。原則的には「昏聲」の字には「昏」，「舌聲」の字には「舌」をその部分字形とすれば良い。この原則をここでは

21) 「𦵏，古文陸字也。」では、「賣」字の説解より「古文陸字」とすべきである。また，SATの畫像では「𦵏」に異體字が使われているが，ここではその詳細について議論しない。例えば，四篇上7b 目部「陸」字の段注を参照していただきたい。

22) 『慧琳音義』のような版本，寫本については校訂や翻刻について議論すべき様々な事柄がある。ここではあくまで『説文』との関係やその解釋の例として提出した。

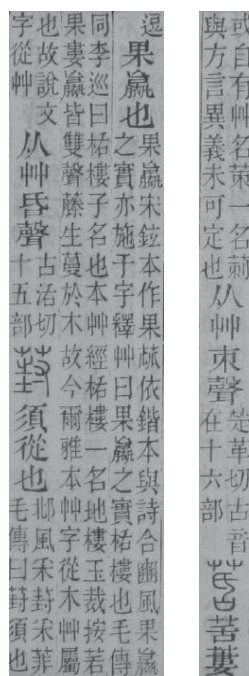


圖6 「菴」(一篇下21b)

「聲符部分字形の原則」と呼ぶことにする。しかし、段氏の用字を見ると必ずしもそのような原則に従わず、通例のように「昏聲」の字に「舌」を用いることも多い。更に艸部「菴」の注で見たように、「昏聲」の字に部分字形「ノ右」を使用することもある。

部分字形として「昏」「舌」「ノ右」をどのように使い分けているのか、小節を改めて議論しよう。

5.4 「栝」と「栝」

ここでは部分字形としての「昏」「舌」「ノ右」の使い分けと、その轉寫について議論しよう。

「昏聲」の字は本来、説文の「昏」の字形で轉寫すべきであるが、「昏聲」の字と「舌聲」の字が別字でないかぎり、段注の各テキストで採用された字形を轉寫に使用しても問題は発生しない。例えば「活」を「活」で轉寫する場合がそれである。但し前小節で述べたように「栝」と「栝」は別字であり、「栝」の代わりに「栝」を使うのは許されない。

段氏が部分字形「ノ右」を使用する字は「昏聲」の字だけである。經典の引用や熟れた語に對して、通常は「舌」の部分字體が用いられている場合、「昏聲」であることを明示するために部分字形「ノ右」を使用する。例えば「栝樹」「栝樓」「栝柏」のような名

稱を、「聲符部分字形の原則」に従って「栢樹」「栢樓」「栢柏」とするのは不自然である。上の艸部「苦」の注にも「栢樓」が出てきたが、これが經典などの引用であれば尙のことである。そこで段注では〔ノ右〕の字形が使われたと想像できる。しかし、この原則がいつも守られる譯ではなく、例えば段注の六篇上 53a 木部「栢」では「栢」字が使われている。

また、段注の標點本〔4〕の艸部「苦」の注でも「栢樓」を「栢樓」に作る。これを一概に誤りとして処理して良いであろうか。他にも段注の六篇上 53a 木部「栢」でも「栢」字が使われている。



(栢) 櫟也。從木昏聲。古活切。十五部。一曰矢栢。櫟弦處。

櫟各本作栢。不可通。今正。釋名曰。矢末曰栢。栢，會也。與弦會也。云櫟弦處者，弦可隱其間也。此亦櫟栢之一端耳。而別言之者。俗但知栢爲矢栢字。嫌義不備。故箸之也。矢栢字經傳多用栢。他書亦用栢。(六篇上 53a)

注に見える「栢」字は説解で「矢栢」と述べられていることや、「會」と疊韻であることから、明らかに「栢」である。この例だけで断定はできないが、段氏もこのように「栢」を使用するのであれば、「栢」を表す意味で「栢」を使用することもあり得る。



(栢) 炊竈木。今俗語云竈栢是也。廣韵云。栢，火杖。栢栢古今字也。從

木。舌聲。臣鉉等曰。當從栢省乃得聲。按徐說非也。栢栢栢等字皆從栢聲。丙見谷部。轉寫譌爲舌耳。他念切。七部。(六篇上 53b)

段氏が注で用いている文字をどのような文字概念であると解釋するかが、轉寫のポイントになる。この状況を少し整理しよう。『説文』では「栢」は「火搔き棒」「たきぎ」を意味し、音は古音七部である²³⁾。また「栢」は「ゆだめ」「やはず」を意味し、音は十五部である。『説文』では「栢」と「栢」は別字で音義ともに重ならないので、この二つの字を示すために、これ以降〈栢〉と〈栢〉を用いる。勿論、ここでは段氏が『説文』の文字として同定する文字概念を表わす。

段注の地の文などで使用されている文字は段氏が通常、論述などに使用している文字概念を表わしている。これらを《栢》《栢》で表現しよう。特に《栢》は〈栢〉を表現す

23) 「栢」の注で「舌聲」に對する疑義が持たれているが、ここでは取り上げない。

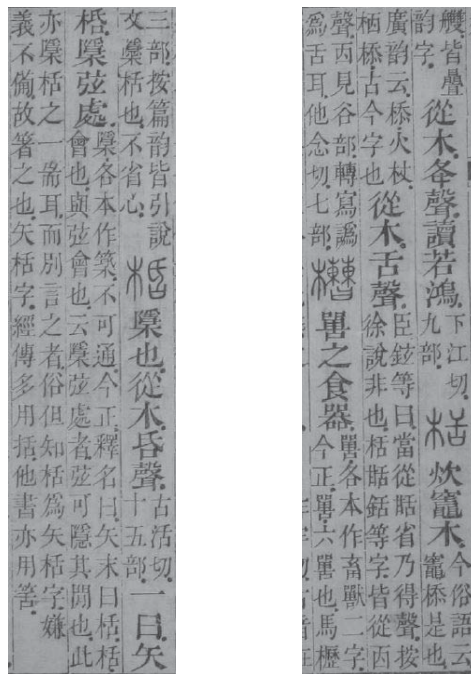


圖7 「栝」(六篇上 53a), 「栝」(六篇上 53b)

るとともに、〈栝〉をも表現し、また、段氏が「栝」字で表現しようとしている他の文字概念もそこには含まれている。

その他に引用文献で使用されている文字を表現する文字概念がある。これには文献の段氏の解釈が含まれ、段氏が作成した釋文という意味合いのものである。特に引用文献に「栝」字が用いられていても、〈栝〉であると解釋した場合には〔ノ右〕の字形が用いられる。

段注で使用されている文字「栝」「栝」に對して、段氏が用いた可能性のある文字概念の候補は以下のように纏めることができる。

1. 説文の文字概念：〈栝〉〈栝〉
2. 引用文献の文字概念（段氏の釋文）
3. 段氏の使用する文字概念：《栝》《栝》

このような文字概念の區別を段氏が用いていたかは、今のところ筆者には分らない。文献學的には否定されるような想定であるかもしれない。

また我々が通常使っている文字や文字コードの文字概念、『廣韻』や様々な字書の文字

概念など、我々が扱うべき文字概念は多様である²⁴⁾。以上が理論的には考え得る文字概念の階層であり、文字が出現する環境に應じて、電子テキスト自身が文字概念を指示できることが将来の機構として必要になるであろう。

6 形式化について

ここまで Unicode 文字集合による段注の轉寫の例を見てきた。本節ではこれらの事例で得た問題点を文献 [11] で議論した形式化という視点から検討しなおそう。特に音注から得られる情報の扱いに注目する。

文献 [11] では Institution 理論を應用して文字體系の形式化を與えた²⁵⁾。文字體系とは文字を中心に構成される論理的な體系で、文字についての事実を記述するために導入される。文字體系はその體系を得るための典據、出所ごとに別々に構成される。例えば、古典籍の場合、釋文を作る際に得られた文字の構造的な體系（文字の張り合いに関する事実）がそれである。また、段注のように文字に関する様々な知識を提供してくれる文献においては、六書に關係した文字に関する事実や、古音分部、それと中古音の韻との關連などが作る體系が文字體系を提供する。

ここでは文字體系と呼んでいるが、必ずしも文字だけを中心に扱う譯ではない。文字概念を周邊から支える體系、『六書音均表』の表一、表二のような音韻情報だけを扱ったものも含まれる²⁶⁾。段注における音注の情報を記述するためには、段注だけを扱うだけではなく、段注から構成される文字體系に『音均表一』『音均表二』『廣韻』が作る三つの體系を〈統合〉して作られる文字體系を利用する²⁷⁾。以下の小節「古音分部の導出の例」と「複数の體系の重ね合わせ」を通して、「文字體系の統合」の具体的な姿について

24) 『漢語大字典』の参照は、「栝」21200.030, 「楮」21218.020 であり, 「栝」は〈栝〉と〈楮〉を含んでいる。

25) Institution 理論は従来、プログラミング言語の意味論や仕様記述などを理論的に扱うために導入されたが、近年では更に廣い範圍の形式的體系を圏論的に記述するために利用されている。ここでは特にモジュール性の記述、オントロジーの意味論的な統合を基礎付ける枠組みに注目している。Institution 理論については文献 [11] であげた参考文献を参照していただきたい。

26) 『六書音均表』『今韵古分十七部表』は中古音の韻目の古音分部への所屬を記述した表であり, 「古十七部諧聲表」は諧聲符の古音分部への所屬を記述した表である。以下、それぞれ『音均表一』『音均表二』と記す。文献 [11] ではこれら二つの表を使って、以下で述べる音韻に関する推論を自動化している。

27) 文献 [11] では〈統合〉という概念を形式化するために Institution という理論を必要とした。「文字體系の統合による漢字情報の形式化」と題した所以である。

見ていこう。

6.1 古音分部の導出の例

漢籍の轉寫においては音注で提供される情報が屢々役に立つことは、本稿の事例でも度々目にする事ができた。段注の翻刻においては「段玉裁の音理」が重要な役割を果たす。まず上古音と中古音の関連に注目して形式化の意味について検討しよう²⁸⁾。

圖 8 は上古音の分部の導出の例である²⁹⁾。各矢印は文獻が提示している音韻的な記述である。例えば、矢印 $Char \xrightarrow{\text{段注古音 (a)}} RGroup$ が示しているのは、文字の集合 $Char$ と古音分部の集合 $RGroup$ の関係、即ちある文字 $c \in Char$ の古音分部が $g \in RGroup$ であるという事實在段注に書かれていることを意味する。

以下はある文字の古音分部を導出するための方法の例である。

1. 段注の古音分部の記述 (a) を通して、文字 $c \in Char$ の古音分部 $g \in RGroup$ が分かる。
2. 『廣韻』の記述 (b) を通して、文字 $c \in Char$ の中古韻 $r \in Rime$ が分かり、『音均表一』の記述 (c) から古音分部 $g \in RGroup$ が分かる。即ち文字 c の古音分部が g であることが分かる。
3. 段注の反切についての記述 (d) を通して、文字 $c \in Char$ の反切下字 $l \in LChar$ が分

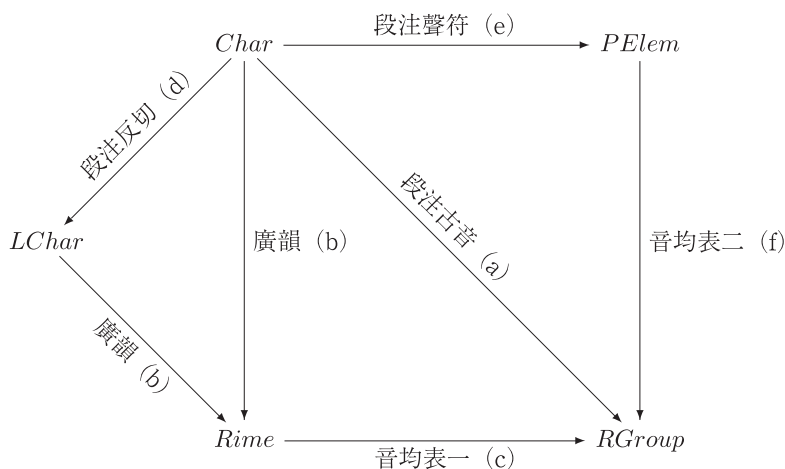


圖 8 古音分部の導出

28) 「段玉裁の音理」については、例えば文獻 [15] を参照していただきたい。

29) 圖 8 は圏論における可換圖ではない。また、矢印は關數ではなく、一般の關係を示している。

り、『廣韻』の記述 (b) を通して、反切下字 $l \in LChar$ の中古韻 $r \in Rime$ が分かる。

これに『音均表一』の記述 (c) を適用して古音分部 $g \in RGroup$ が分かる。

4. 段注の記述 (e) を通して、文字 $c \in Char$ の聲符 $p \in P\ Elem$ が判明する。『音均表二』の記述 (f) から聲符 $p \in P\ Elem$ の古音分部 $RGroup$ が分かる。この聲符 p の古音分部 g を対象となった文字 c の古音分部とすることができる。
5. 段注の記述 (e) で得られた聲符 $p \in P\ Elem$ は文字の集合 $Char$ に含まれる。聲符を上記の四つの導出法における文字 $Char$ として扱うことによって、更なる古音分部の導出を行なうことができる。例えば、聲符は更に基本的な聲符に歸着できる場合（段注の記述 (e) に聲符 $p \in Char$ の更なる聲符 $p' \in P\ Elem$ が記述されている場合）があり、その聲符 p' の古音分部 g を段注の記述 (a) から得ることができる。

ある文字の古音の分部（導出1）と今韻の對應する分部（導出3）が異なるとき、段注では音の不規則變化が起きたと考え、「古音在幾部」と記述される。例えば、五篇上可部の「奇」は導出1では十七部であるが、導出3では十六部であり、「古音在十七部」と記されている。この場合、十七部から十六部へと音韻變化が起こったと考える。

古音分部の導出についてのみ述べたが、一般に文字についての様々な性質を導出することが要求される。その際に性質の提示が文献の記述であるのか、記述からの推論であるのかを區別される必要がある。文字概念の電子的な表現にはこのような推論を容易にする形式化が望まれる。

6.2 複数の體系の重ね合わせ

上で述べた古音分部の導出には複数の文字體系の重ね合わせが含意されている。體系の重ね合わせによって様々な推論が可能になるのである。圖8の各ノードは複数の文献に含まれる「対象」が重ね合わされている。圖の矢印は各文献の記述を表現している譯だが、異なる文献からの矢印がノードを共有しているところでは、各文献からの「対象」に重ね合わせが起こっている³⁰⁾。

文字 $Char$ からは段注の矢印 $Char \xrightarrow{\text{段注古音 (a)}} RGroup$ と『廣韻』の矢印 $Char \xrightarrow{\text{廣韻 (b)}} Rime$ の二つの矢印が出ているので、文字 $Char$ のところで段注と『廣韻』で記述されている文字の重ね合わせがある。また、韻 $Rime$ には『廣韻』の矢印 $Char \xrightarrow{\text{廣韻 (b)}} Rime$ が入り、『音均表一』の矢印 $Rime \xrightarrow{\text{音均表一 (c)}} RGroup$ が出ているので、韻 $Rime$ のところで

30) 実際には「対象」だけでなく「性質」についての重ね合わせも行なう必要があるが、ここでは詳しく述べない。文献 [11] を参照していただきたい。

『廣韻』と『音均表一』で記述されている韻が重ね合わせられている³¹⁾。

重ね合わせが起こるのは、『廣韻』と段注で文字の張り合いが異なるといった場合だけでなく、例え同じ音韻體系（今韻二百六韻）が使われる場合でも異なる文獻に含まれる韻の間では起こっている³²⁾。つまり重ね合わせの「対象」には文字だけでなく、文字に関連する周囲の対象、韻や古音分部なども含まれる。

以上、文獻から作られる個々の文字體系の形式化だけではなく、それらを利用するためには複数の文字體系の相互の関係も形式化の対象となっている。即ち、複数の文字體系の重ね合わせの仕組みを実現することが要請される。文獻[11]の用語では文字體系の〈統合〉によってこれを實現した。

6.3 形式化への要求

電子的な翻刻では文字概念の形式化を何如にするかということが問題となった。漢字を定義するために、その「形音義」を形式化することの難しさは、これまでの研究史が明かにしてきたことではないだろうか。UniHanの文字概念のようにpropertyとその値の組で与えられるような文字概念では、電子化の際に結局、標準字書のような他文獻の文字概念を援用することになってしまう。

文獻[11]では、漢字をその「形音義」によって直接形式化するのではなく、対象となる文獻における文字の張り合いという構造的な體系、及びその文字體系を取り巻く様々な體系との関係を記述することによって、文字概念を記述する試みを行なった。文字の周囲の體系として、現在は音韻を中心に扱っているが、今後は引用による他文獻との文字の関連、統語論的な體系との関連、視覚的な表現などが形式的な體系の対象となる。

これまでの段注の電子的な翻刻で見てきたように、漢籍の電子化に二種の階層（ある意味で深層と表層）が存在することが見てとれた。一つは文獻の解釋を通して構成された釋文の文字概念であり、これは上で見た通り文字概念の形式化の問題である。もう一つの層は文獻の表層としてのテキストに文獻構造に對應する多様性があることであり、文獻構造によって指示する文字概念が異なるといった問題である。例えば、小節5.4で見たように、段注では『説文』、段氏の論述、引用文獻の釋文などに使われる様々な文字

31) ここでの「対象」の記述には『廣韻』・段注・釋文などの文字概念の確定とそれへの「名指し」が終了していると假定している。

32) 韻の重ね合わせの例としては、例えば二つの文獻の韻に異なる文字が使われている場合、『廣韻』の下平聲二十一侵は『音均表一』では「侵」であり、上聲四十七寢は「寢」のような場合がある。

が指示されていて、それをテキストとして表現するための工夫が見られる。

将来の電子テキストの形式化には、この二つの層を分けて考えるのが良いであろう。

1. 文字體系の構築と複数の文字體系の統合、これには以下が含まれる。
 - ・ 漢字の性質：文献ごとの可能世界的な表現
 - ・ 推論による整合性の検査
2. テキストの表層を表現する機構

文献 [11] は、文字體系の構築と複数の文字體系の統合を形式化したものである。テキストの表層を表現する機構については今後の課題とする。

7 お わ り に

本稿で述べた翻刻では文字の表現に新しい方法を用いず、敢えて Unicode 文字集合という広く普及している方法を用いた。Unicode 文字集合による翻刻の限界や問題点を浮き彫りにすることができたのではないだろうか³³⁾。

また、電子的な翻刻における文字コードの指定という事象が形や見た目だけの再現では済まず、解釈や意味まで指定していることを事例において十分に示せたと思う。ここでの段注の電子的な翻刻は、文字のモデルとその表現の仕組みに対する第一着手であり³⁴⁾、今後節 6.3 で述べた形式化への要求を精密化することが課題である。

文献 [10] では、複数の古辭書の情報を提供する Web サイトととしてプロトタイプを実装した³⁵⁾。『説文解字』大徐本・『段注』・『廣韻』・『漢語大字典』である。プロトタイプにはここでの轉寫データは反映されていないが、将来的に本稿の結果を取り入れる予定である。

33) 本稿の目的は飽く迄、情報學の視點から古漢籍の翻刻の事例を述べることであり、文献學の分野からは不適切であるとお叱りを受ける點も多々あると思う。ご意見ならびにご教示願えれば幸いである。

34) 本稿は文字表現の電子的な姿を段注の翻刻という事例から眺めたものであり、部分字形への言及等、文字の視覺的な表現に『説文』の規範を持ち込むという意圖はない。かえって、視覺的な表現も含めて、文字の體系性と歴史的な關連を形式化するために、文献 [11] のような枠組みを必要としていることを強調するものである。

35) Han Morphism System <http://www2.dhii.jp:3000/>, 本サイトは一般財團法人 人文情報學研究所の協力により公開されている。

■ 謝辭 Christian Wittern 先生には京都大學人文科學研究所の研究班「人文情報學の基礎研究」に参加を認めて下さり、多くの有益な助言と機會を與えて戴きました。「點注會」参加の皆さま、特に古勝隆一先生からは様々なことを學ばせて戴いております。Han Morphism System の公開にあたり、財人文情報學研究所の助力を得ました。多くの方々から感謝し盡くせない援助を戴きました。最後にいつも支えてくれる妻留美と娘に感謝します。

附錄 A 「字畫原始」の例

ここでは部分字形の参考のために「字畫原始」の一部の轉寫を示した。「字畫原始」とは山梨稻川³⁶⁾の『文緯』[14]に見えるもので、『説文』における字形の元となる基本形を列舉したものである。

丨	丨	U+4E28	一篇上丨部
乚	ノ	U+4E3F	十二篇下ノ部
ㄣ	ㄣ	U+4E40	十二篇下ノ部
ㄥ	ㄥ	U+4E85	十二篇下ノ部
ㄨ	ㄨ	U+2010C	十二篇下ノ部
ㄩ	ㄩ	U+20086	十二篇下ノ部
ㄴ	ㄴ	U+4E41	十二篇下ノ部
ㄷ	ㄷ	U+5382	九篇下ノ部
ㄹ	ㄹ	U+200CA	十二篇下ノ部
ㄺ	ㄺ	U+21FE8	十一篇下ノ部
ㄻ	ㄻ	U+5DDC	十一篇下ノ部
ㄽ	ㄽ	U+51F5	二篇上ノ部
ㄿ	ㄿ	U+20674	五篇上ノ部

附錄 B 「賣」と「賣」

B.1 「賣」「賣」の異體字

この二字に關係する Unicode の異體字を列舉する。切り分けは『漢語大字典』に據る。

36) 江戸時代の説文學者。明和 8 年 (1771 年)―文政 9 年 (1826 年)。

各字にそえた文献は『漢語大字典』が根拠としている文献である。これらの例を見る限り、異體字関係を字體だけで見分けるのは難しいと思われる。

六篇下 2b 出部「賣」莫遯切

1. 「賣」 U+8CE3 『説文』
2. 「賣」 U+27DD3 『説文』『玉篇』
3. 「賣」 U+27DF5 『説文』『正字通』
4. 「賣」 U+27DD7 『篇海類編』
5. 「賣」 U+27DE8 『字彙補』
6. 「賣」 U+27DF7 『九經文字』

六篇下 21b 貝部「賣」余六切

1. 「賣」 U+27DA0 『説文』
2. 「賣」 U+27E07 『説文』
3. 「賣」 U+27DCF 『字彙』『正字通』

『説文』で六篇下貝部「賣」余六切 𧮮 に従う字は以下である。() 中は重文, [𧮮] は新附字 (八篇下見部) である。

𧮮 𧮮 𧮮 𧮮 𧮮 𧮮 𧮮 (𧮮) 𧮮 𧮮 𧮮 𧮮 𧮮 𧮮 [𧮮] 𧮮 𧮮 𧮮 𧮮

これを標準字形 (Unihan の画像) で確認すると, Unicode の標準字形は『説文』の造字の状況を反映していない。以下にフォント HanNom による表示を示す。

𧮮 𧮮 𧮮 𧮮 𧮮 𧮮 (𧮮) 𧮮 𧮮 𧮮 𧮮 𧮮 [𧮮] 𧮮 𧮮 𧮮 𧮮 (𧮮)
(HanNom)

またフォント花園明朝による表示を以下に示す。HanNom に比較して, 部分字形として「賣」を使用するものが多いことが分かる。このようにフォントに依存した字形があるため, Unicode を使用した転寫では注意が必要である。

𧮮 𧮮 𧮮 𧮮 𧮮 𧮮 (𧮮) 𧮮 𧮮 𧮮 𧮮 𧮮 [𧮮] 𧮮 𧮮 𧮮 𧮮 (𧮮)
(花園明朝)

B.2 貝部「賣」の使用例

貝部「賣」（余六切）を聲符とする例を見る。各項目の括弧内は、大徐本と段注の聲符の記述で、誤りがある場合に記號×を付けた。

1. 三篇上言部「讀」誦書也。从言賣聲。(説文「賣」× 段注「賣」)
2. 六篇上木部「櫝」匱也。从木賣聲。一曰木名。又曰：大椀也。(説文「賣」 段注「賣」×)
3. 七篇上片部「牘」書版也。从片賣聲。(説文「賣」 段注「賣」×)
4. 十篇上黑部「黷」握持垢也。从黑賣聲。易曰：再三黷。(説文「賣」× 段注「賣」)
5. 十一篇上水部「瀆」溝也。从水賣聲。一曰邑中溝。(説文「賣」× 段注「賣」)
6. 十二篇下女部「嬪」媒嬪也。从女賣聲。(説文「賣」× 段注「賣」×)
7. 十二篇下匚部「匱」匱也。从匚賣聲。(説文「賣」× 段注「賣」)
8. 十三篇上糸部「續」連也。从糸賣聲。(説文「賣」× 段注「賣」)
9. 十四篇下自部「隕」通溝也。从自賣聲。讀若瀆。(説文「賣」× 段注「賣」)
10. 四篇下夕部「殯」胎敗也。从夕賣聲。(説文「賣」× 段注「賣」×)
11. 一篇下艸部「蕢」水鳥也。从艸賣聲。詩曰：言采其蕢。(説文「賣」 段注「賣」)
12. 六篇下貝部「贖」質也。从貝賣聲。(説文「賣」 段注「賣」)
13. 八篇下見部「覲」見也。从見賣聲。(説文新附「賣」)
14. 二篇下彡部「遺」媒遺也。从彡賣聲。(説文「賣」 段注「賣」)
15. 三篇上誥部「讎」痛怨也。从誥賣聲。春秋傳曰：民無怨讎。(説文「賣」 段注「賣」)
16. 三篇下革部「鞮」弓矢鞮也。从革賣聲。(説文「賣」 段注「賣」)
17. 八篇上人部「價」賣也。从人賣聲。(説文「賣」 段注「賣」×)

B.3 「賣」「賈」に関する異同

以下は出部「賣」と貝部「賈」に関連して、『説文』と段注で異なる字形を使用している例（下線部）である。

1. 二篇下行部「衍」行且賈也。从行从言。(説文「賣」 段注「賣」)
2. 五篇下門部「市」買賈所之也。市有垣，从門从入。入，古文及，象物相及也。之省聲。(説文「賣」 段注「賣 a」)
3. 六篇下貝部「賈」賈市也。从貝而聲。一曰坐賣售也。(説文「賣」 段注「賣 a」)
4. 六篇下貝部「販」買賤賈貴者。从貝反聲。(説文「賣」 段注「賣 a」)

但し、「賣 a」は横目の中が特殊な形をしている（段注五篇下 26b 冂部「市」を参照）。

附録 C 「昏聲」の字

段注における「昏聲」の字は以下の 18 字あり、その字の古音分部は全て十五部である。

1. 一篇上 13b 示部「𠄎」
2. 一篇下 21b 艸部「𦰩」
3. 二篇下 5a 辵部「𨔵」
4. 二篇下 23a 齒部「𪔐」
5. 三篇上 15b 言部「𦘔」
6. 四篇上 50a 鳥部「𪖇」
7. 四篇下 16b 骨部「𪖇」
8. 四篇下 47b 刀部「刮」
9. 六篇上 53a 木部「𣎵」
10. 七篇上 48b 禾部「𪔐」
11. 八篇上 19a 人部「𠄎」
12. 九篇上 7a 頁部「𦘔」
13. 九篇上 24b 髟部「髻」
14. 十一篇上二 3a 水部「活」
15. 十二篇上 17b 耳部「聒」
16. 十二篇上 46a 手部「括」
17. 十二篇下 16b 女部「𪔐」
18. 十四篇上 25b 金部「銛」

また、「活聲」「銛聲」とする字がそれぞれ一つずつある。

1. 十二篇上 15a 門部「闕」（活聲）
2. 十篇下 41a 心部「𦘔」（銛聲）

参 考 文 献

[1] 許力以主任，徐中舒主編：『漢語大字典』，四川辭書出版社，湖北辭書出版社，1986-1990.

- [2] 許慎撰, 徐鉉校定:『說文解字』, 中華書局, 1963.
- [3] 許慎撰, 段玉裁注:『說文解字注』, 影印, 上海古籍出版社, 1981.
- [4] 許慎撰, 段玉裁注, 許惟賢整理:『說文解字注』, 鳳凰出版社, 2007.
- [5] R. S. Cook:《說文解字・電子版》 Shuo Wen Jie Zi-Dianzi Ban : Digital Recension of the Eastern Han Chinese Grammaticon, PhD Dissertation, Department of Linguistics, University of California, Berkeley, 2003.
- [6] J. D. Allen, et al. eds.: The Unicode Standard, Version 5.0, The Unicode Consortium, 2006.
- [7] M. Fowler: UML Distilled: A Brief Guide to the Standard Object Modeling Language (3rd ed.), Addison-Wesley, 2003.
- [8] J. H. Jenkins, R. Cook, eds.: Unicode Han Database (Unihan), Unicode Standard Annex 38, [http :
//www.unicode.org/reports/tr38/](http://www.unicode.org/reports/tr38/), 2006-2010.
- [9] 「常用漢字表」, 平成 22 年内閣告示第 2 號, 2010.
- [10] 白須裕之: 古辭書のテキストアーカイブズ構築について —— 文字轉寫の理論とその應用, 情報處理學會「人文科學とコンピュータシンポジウム」論文集, 頁 395-402, 2011.
- [11] 白須裕之: 文字體系の統合による漢字情報の形式化 ——『說文解字注』における音注を事例として, 情報處理學會「人文科學とコンピュータシンポジウム」論文集, 頁 209-216, 2013.
- [12] 須永哲矢, 堤智昭, 高田智和: 明治前期雜誌の異體漢字と文字コード ——『明六雜誌』を事例として ——, 情報處理學會「人文科學とコンピュータシンポジウム」論文集, 頁 381-388, 2011.
- [13] 豐島正之: 「JIS に無い字」をめぐって, 月刊しにか 3(2): 16-22, 大修館書店, 1992.
- [14] 山梨稻川: 文緯「字畫原始」, 所收 山梨稻川集刊行會編: 山梨稻川集第 1 卷, 山梨稻川集刊行會, 1929.
- [15] 賴惟勤監修, 說文會編: 說文入門, 大修館書店, 1983.